DOCUMENT RESUME

ED 093 933 TM 003 751

AUTHOR Halinski, Ronald S.; Tcheng, Tse-Kia
TITLE Systematic Student Input into Evaluation of an Educational Innovation Program.
PUB DATE Apr 74
NOTE 11p.; Paper presented at the Annual Meeting of the American Educational Research Association (59th, Chicago, Illinois, April 1974)

EDRS PRICE MF-$0.75 HC-$1.50 PLUS POSTAGE
DESCRIPTORS Curriculum Evaluation; Curriculum Planning; *Educational Innovation; Higher Education; *Instructional Improvement; *Program Effectiveness; *Program Evaluation; *Student Participation

ABSTRACT

Under a program of educational innovation implemented by the university, some 20 percent of the faculty were engaged in activities designed to improve the quality of instruction. Financial support exceeded one-half million dollars. The purpose of this study was to determine the program's overall impact. Baseline data was gathered in March 1973, and the survey was replicated one year later. Included among the data gathered from students was: (1) expectations and actual progress toward the attainment of selected educational objectives, (2) activities and interest in the general culture, (3) learning styles, and (4) general satisfaction with various instructional delivery systems. Each survey included some 4,000 respondents. (Author)

ED 093933

T' 003 251

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

SYSTEMATIC STUDENT INPUT INTO EVALUATION OF AN EDUCATIONAL INNOVATION PROGRAM

Ronald S. Halinski and Tse-Kia Tcheng
Illinois State University

INTRODUCTION

The ongoing purpose of this project is to develop a longitudinal data base which would provide for systematic student input into the curricular planning and evaluation functions of the university. An immediate need for the data base was brought about by the implementation of a program of educational and instructional innovation. Under that program 35 projects were funded and each was separately evaluated. These projects involved most of the departments of the university and included as a major thrust either development, implementation, or evaluation type activities. (For example one project restructured three courses into a single competency based course; another project used professional actors to present live dramatizations in regular classes; a third engaged in an extensive evaluation of the department's principal general education course.) In addition, 50 instructors received small grants to develop innovative instructional practices. Totally, 20 per cent of the faculty were engaged in activities designed to improve the quality of instruction. Financial support exceeded one-half million dollars. The objectives, activities and evaluation of these projects are described in more detail elsewhere.[1]

In addition to any successes enjoyed by the individual projects one might also expect a type of ripple effect. Project faculty taught additional courses and the possibility of carryover existed. Project ideas provided examples for others to emulate. But probably most important the formalized innovation program provided an explicit commitment to improving the quality of instruction and legitimized such endeavors as appropriate for professional activities. Thus what appeared necessary was the development of a strategy which might get at the overall impact of the program.

PROCEDURE

Baseline data was gathered in March, 1973. Four forms of a modularized questionnaire were developed from measures in the Higher Education Measurement and Evaluation Kit prepared by the staff of the Higher Education Evaluation Program under the Directorship of C. Robert Pace. Information of the following types was elicited: (1) overall satisfaction with the instructional program; (2) general satisfaction with various instructional formats; (3) expected and actual progress toward the attainment of broad educational objectives and benefits classified into vocational, general education, critical thinking and human relations categories; (4) activities and interest in the general culture; (5) student characteristics. 232 items were divided into subgroups and these were distributed among the four questionnaires so that no form would take longer than 15 minutes to complete. All undergraduate classes meeting at 9 AM on a Wednesday were included in the survey. Questionnaires were completed anonymously during the

class period. To increase the likelihood of representativeness the four forms were intermingled sequentially when they were packaged for distribution. Some 4000 responses were gathered for an 80% return. The nonrespondents were students who were absent or who were in classes where the instructor declined to take part. Neither source of bias was considered serious. In particular it appeared that the subgroup of students who miss a proportion of their classes were actually represented since some of the students attending the 9 AM class would likely miss classes at other times during the day.

The original survey with two modifications was replicated in mid-Frebruary, 1974. (The difference in time of year was due to a change in the school calendar.) The two modifications included eliminating some subgroups of items and distributing the remaining ones among three forms of the questionnaire and administering the questionnaire to 10 AM classes. Some 4600 responses were received for an approximate return of 75 per cent.

RESULTS

Overall Satisfaction. If students feel that what they are doing is worthwhile and view the university favorably, they are more likely to be productive and to take advantage of the many opportunities available to them. Table 1 summarizes the results of several attempts to determine globally student feelings regarding the university and the quality of education they are receiving. Class means were weighted according to the proportion of students in each class based on actual headcount. The proportions changed from '73 to '74 but in each case were in the neighborhood of .25. To produce a measure for comparison purposes which is independent of the changing pattern of enrollment, the class means were also weighted equally to arrive at a composite value for the four classes.

In general, the results show a movement to greater satisfaction with the university although the movement is small. As to be expected, the large sample sizes allow for small observed differences to be statistically significant. As yet, we do not have sufficient experience with these measures to make judgments of practical significance. There are two other observations based on these tables which you may have noted: first, the results are most favorable for the senior class and secondly, the method of weighting did not make any material difference in the results.

Informational Format. Students were asked to check which types of instructional formats they had experienced at ISU and their overall degree of satisfaction with such experience (Scale: 1 = Highly Dissatisfied to 5 = Highly Satisfied). In Table 2 two types of data are reported for both 1973 and 1974. First, the percentage of students who indicated they had experienced an instructional format at least once at the university is shown. Secondly, of those who had experienced the particular format, the mean "satisfaction" ratings are reported.

Table 1. General Satisfaction of Students With the University.

Item	Class	1973 Mean	1974 Mean	Significance
1. Overall Quality of Instruction (Scale: 1 = Excellent, 2 = Good, 3 = Fair, 4 = Poor)	Fr	2.16	2.17	n.s.
	Soph	2.21	2.21	n.s.
	Jr	2.20	2.16	n.s.
	Sr	2.25	2.17*	$p < .05$
Weighted Mean (Headcount)		2.21	2.18*	$p < .05$
Weighted Mean (Weight = .25)		2.21	2.18*	$p < .10$
2. Relevance of Educational Experiences (Scale: 1 = Definitely Yes, 2 = Generally Yes, 3 = Generally No, 4 = Definitely No)	Fr	2.14	2.14	n.s.
	Soph	2.20	2.15	n.s.
	Jr	2.11	2.05*	$p < .10$
	Sr	2.18	2.08*	$p < .01$
Weighted Mean (Headcount)		2.16	2.10*	$p < .001$
Weighted Mean (Weight = .25)		2.16	2.10*	$p < .001$
3. University's Concern for the Individual (Scale: 1 = Definitely Yes, 2 = Generally Yes, 3 = Generally No, 4 = Definitely No)	Fr	2.68*	2.76	$p < .05$
	Soph	2.80	2.84	n.s.
	Jr	2.78	2.76	n.s.
	Sr	2.85	2.76*	$p < .05$
Weighted Mean (Headcount)		2.78	2.78	n.s.
Weighted Mean (Weight = .25)		2.78	2.78	n.s.

*Indicates more favorable response when difference is significant.

Approximate sample sizes for each question:

	1973	1974
Fr	N=1160	N= 827
Soph	N= 955	N= 970
Jr	N= 980	N=1450
Sr	N= 720	N=1140

Table 2. Student Experience and Satisfaction with Various Instructional Formats

Instructional Format	Per Cent of Students Experiencing Instr. Format[1]		Mean Satisfaction Rating[2]		Sig.
	1973	1974	1973	1974	
Large Lecture Class	99	98	2.80	2.94*	$p < .001$
Small Class with Instructor-Led Discussions	97	95	3.91	3.96	n.s.
Individual Research as Part of Course Work	89	90	3.53	3.57	n.s.
Group Projects as Part of Course Work	82	81	3.08	3.20*	$p < .01$
Lecture Class with Scheduled Discussion Sections	81	81	3.14	3.33*	$p < .001$
A Laboratory Course	75	65	3.30	3.45*	$p < .001$
Student-Led Discussion Groups	74	71	3.29	3.38*	$p < .05$
Team Teaching	70	65	3.40*	3.26	$p < .01$
Group Research as Part of Course Work	67	67	3.01	3.11*	$p < .05$
Video-Taped Lectures	51	48	2.32	2.55*	$p < .001$
Courses Involving Community Experience	38	36	3.49	3.68*	$p < .01$
Self-Instructional Packages in Learning Laboratories	33	34	2.79	2.86	n.s.
Independent Study	30	31	3.39	3.49	n.s.
Part or All of Course Work Conducted Off Campus	26	32	3.28	3.42*	$p < .05$
Residence Hall courses	19	18	2.94	3.33*	$p < .05$

[1]For 1973, n=1890; for 1974, n=1685. Satisfaction ratings were completed only by those students who experienced the format at the university.

[2]Scale: 1=Highly Dissatisfied, 2=Dissatisfied, 3=Neutral, 4=Satisfied, 5=Highly Satisfied. *indicates higher mean satisfaction rating.

NOTE: Class means were weighted equally to arrive at the overall mean.

We make the following observations:

1. The most widely experienced instructional formats were the large lecture type class and the small class with instructor led discussions.

2. The percentage of students experiencing a particular format did not change to any large extent from '73 to '74 with the exception that there were 10 per cent fewer students who had experienced a laboratory course in 1974.

3. The small class with instructor-led discussions received the highest satisfaction rating on both years; the difference between the two years was not statistically significant.

4. Of the 15 instructional formats, 10 received significantly higher satisfaction ratings in '74.

The development of mediated instructional materials and their implementation received heavy emphasis in the innovation program. Because the developmental effort required is extensive it may be too early yet to assess the impact. However, the results provide some useful insights for future direction. First of all, the significant increase in mean satisfaction rating for video-tape lectures is promising. Less encouraging is the fact that a similar result did not occur for self-instructional packages in which mediated materials are a key ingredient. Secondly, the relatively low satisfaction ratings for the video-tape lecture and self-instructional formats indicate that the mere production of such mediated materials will not guarantee acceptance by the students. It seems clear that any movement in the direction of technology should be carefully planned and coordinated, adequately funded and fully evaluated. What may appear to be obvious is too often lacking in practice.

Community-based experiences and large lecture classes also received considerable attention among the projects funded and the evidence is favorable: both showed significant increases in student satisfaction. Further, community-based experiences were well-received by the students as suggested by the relatively high mean satisfaction rating. However, with large lecture classes it is a different matter. The relatively low mean rating very likely indicates student tolerance of such classes in general, and the question then becomes whether or not quality education can occur under such circumstances. While the present data does not provide an answer, it does serve to highlight the question.

Educational Benefits. There were 25 statements[2] concerned with educational objectives/benefits associated with college. Students were asked to respond to these statements in terms of their "actual" and "preferred" progress toward attainment on a scale which ranged from 1 (Little or None) to 5 (Very Much). A priority listing for each class was derived by ranking these statements on the basis of the mean "preferred progress" rating. In addition these statements were also ranked on the basis of the mean "actual progress" rating. The difference between the two means for each statement can be viewed as a measure of discrepancy for that particular objective/benefit. Edited versions of these statements appear in Table 4.

The magnitude of the rank-order correlations (Table 3) indicate that no large changes in ranks occured in the 1974 survey for the Preferred and Actual progress dimensions. The somewhat lower correlation between the junior class Discrepancy rankings for the two years was due primarily to the relatively smaller discrepancies for certain of the human relations benefits in 1974. The correlations between the Actual and Preferred ratings show that in general students tend to perceive themselves as making relatively more progress in those areas which are viewed as more important to them.

Table 3. Rank-Order Correlations of Mean Ratings for the 25 Statements of Selected Educational Benefits by class

	Fr.	Soph.	Jr.	Sr.
Preferred progress ratings: 1973 vs. 1974	.91	.91	.97	.90
Actual progress ratings: 1973 vs. 1974	.96	.98	.91	.95
Discrepancy ratings: 1973 vs. 1974	.92	.94	.76	.88
1973: Preferred vs. Actual ratings	.61	.51	.66	.67
1974: Preferred vs. Actual ratings	.72	.62	.71	.71

NOTE: For n = 25, $p < .01$ for $r \geq .51$ (two-tailed test)

The mean of the rank-order correlations among the four classes for the Preferred, Actual and Discrepancy rankings were .89, .93 and .88. Because of this the means for each class were weighted equally to arrive at an overall mean for each of the three dimensions for both 1973 and 1974. The ranks of these means are reported in Table 4.

Table 4. Ranks of Educational Benefit Statements for Preferred Progress, Actual Progress and Discrepancy

Statements by Category	Preferred Progress 1973 Rank	Preferred Progress 1974 Rank	Actual Progress 1973 Rank	Actual Progress 1974 Rank	Discrepancy 1973 Rank	Discrepancy 1974 Rank
Human Relations						
Development of an identity	1	1	8	6	3	3
Social development	2	2	2	2	17	17
Personal development	4	3	4	3	16	20
Tolerance of others	3	4	1	1	23	23
Development of a personal philosophy	17	16	15	14	13	14
Vocational						
Background for further education	7	6	10	11	9	8
Vocational training	9	9	25	25	1	1
Discovery of vocational intersts	15	19	21	21	4	6
Critical Thinking--to develop						
Open-mindedness	5	5	3	4	20	19
Intellectual curiosity	6	7	9	9	10	11
Ability to select appropriate information	8	8	13	13	8	9
Intellectual honesty	13	11	12	12	12	12
Desire for order	16	15	14	15	14	13
Ability to recognize assumptions	20	18	20	16	11	10
Ability to define problems	14	14	18	18	7	5
General Education						
Current issues & problems in society	10	13	19	20	2	4
Develop interests in new fields	12	12	7	7	15	18
Terminology & facts in various fields	11	10	6	8	19	15
Awareness of different cultures	18	20	5	5	24	25
Effective communication	19	17	22	22	5	2
Appreciation of moral & ethical standards	21	21	11	10	25	24
Quantitative thinking	22	22	24	23	6	7
Broadened literary appreciation	23	23	17	17	21	21
Aesthetic sensitivity	24	24	16	19	22	22
Understanding the nature of science	25	25	23	24	18	16

NOTE: Means ranked according to numerical value from High = 1 to Low = 25.

There are several observations which can be made from the table.

1. The Human Relations type benefits such as development of an identity, social development, personal development and tolerance of others ranked highest in priority. Certain benefits of a general college education well accepted by faculty, such as, development of quantitative thinking skills, broadened literary appreciation, aesthetic sensitivity, and understanding the nature of science, ranked lowest in priority.

2. Perceived actual progress was greatest for the following: tolerance of others, social development, personal development, open-mindedness and awareness of different cultures. Perceived actual progress was least for the following: vocational training, understanding the nature of science, quantitative thinking, effective communication and discovery of vocational interests.

3. Among the largest discrepancies were included the following: vocational training, development of an identity, current issues and problems in society, and effective communication. (To place the magnitude of these discrepancies in perspective, the largest value was 1.33 for "vocational training" while the smallest was .59 for "awareness of different cultures". Across all objectives the marked tendency is for students to rate "preferred progress" higher than "actual progress".)

It is possible for gains to occur and have the ranks remain relatively stable. Thus, it might be well to consider the differences in mean ratings between the two surveys. Based on equal weighting of class means to arrive at an overall mean, significant increases ($P < .05$) were observed in Actual progress on the following: (a) Personal development-increase from 3.32 to 3.39, (b) Development of an identity-increase from 3.15 to 3.24, (c) Quantitative thinking-increase from 2.72 to 2.81. There were no significant decreases nor were any of the differences between the mean discrepancies significant.

When the differences in mean discrepancies were tested by class, four were significantly lowered ($p < .05$). However, when you consider 100 tests were run, little, if any, importance can be attached to these. By class, there were 27 differences in "Actual progress" means which were significant ($p < .05$). Ten of these indicated greater progress in the '73 survey and they occurred in total for the freshman class. The remaining seventeen of the significant differences indicated greater progress in the '74 survey and were associated primarily with the junior and senior classes. The majority were in the general education and human relations area. By the nature of the project funded it would be difficult to attribute the gains in the human relations category to the formalized program especially since there is a currently strong movement for establishing local chapters of national social fraternities and sororities.

Our experience in dealing with faculty members regarding evaluation of their particular projects demonstrated to us the difficulty in discussing instructional objectives. Inevitably such attempts would lapse into discussions of process, that is, instructional format. It is our conjecture at this point that the gains exhibited in student satisfaction with instructional format reflect this emphasis and the lack of similar type gains regarding "Actual progress" or decreasing the discrepancy values reflects the lack of emphasis on instructional objectives.

Student Characteristics. The nature and quality of learning a student engages in is reflected to large extent by behavior during lectures and periods of study. The results in Table 5 indicate that students took notes during lectures primarily for tests as opposed to taking notes for their own interests. However, from '73 to '74 there were significant differences in the direction of higher quality learning behaviors. Again, to put the mean values in perspective, in 1974, 73% of the students indicated that "very frequently" they took notes primarily for tests opposed

to 26% who indicated that "very frequently" they took notes primarily for their own interests. Regarding periods of study there were no differences in behavior indicated between '73 and '74 with the exception that students appeared to spend more time thinking about applications of what they were studying.

Table 5. Student Behavior During Lecture Classes and Study Periods

	1973 Mean (N=950)	1974 Mean (N=2630)	Significance
During Lectures:[1]			
Take notes primarily for tests	2.68	2.58*	$p < .001$
Take notes primarily for personal interests	1.96	2.05*	$p < .001$
Relate what instructor says to other things	2.21	2.26*	$p < .05$
During Study Periods:			
Read assignments without understanding them	1.58	1.56	n.s.
Memorize facts	2.11	2.09	n.s.
Relate concepts to personal experience	2.16	2.20	n.s.
Think about applications of the material	2.12	2.22*	$p < .001$

[1]Scale: 1 = Seldom or Never; 2 = Often; 3 = Very Frequently

*Indicates more favorable value

The extent to which individuals engage in certain kinds of activities is a reflection of their interests and attitudes. Eight brief activity scales related to broad general education objectives were selected from the KIT[3] for inclusion in the survey. These scales sampled behaviors which range from commonplace activities to those which require more effort and thus imply a more intensive level of involvement. On each scale students were asked to check those activities they had engaged in during the past year. The score for a given scale was the number of items checked. The results are presented in Table 6. One particular application of these scales which is appealing is their use as a pre- and post-measure in the evaluation of individual projects or courses. The availability of campus-wide norms for appropriate identifiable subgroups could serve as a partial solution to the problem of setting standards in the formation of goals. Such norms might be viewed as minimal standards.

Table 6. Activities and Interest in the General Culture

Scale	Number Of Items	1973 N	1973 Mean	1974 N	1974 Mean	Sig.
Art	7	769	2.98	712	2.96	n.s.
Music	9	931	6.08	830	6.02	n.s.
Literature	8	852	3.84	773	3.76	n.s.
Community Affairs	7	842	2.32	721	2.36	n.s.
Drama	7	902	3.90	819	4.02*	$p < .10$
International & Intercultural Affairs	8	808	3.17*	721	2.95	$p < .01$
Science	10	788	3.18*	677	2.98	$p < .05$
National and State Politics	10	814	3.57	774	4.01*	$p < .001$

*Indicates more favorable value

For students to be deeply involved in academic work and to become independent learners is considered by many faculty to be the essence of a college education. In a speculative vein two related scales from the KIT were included. The first, titled Style of Learning-Academic, is described as measuring "the style of one's effort in relation to the acquisition of knowledge and understanding from courses and readings, such as participation in class discussions, talking with professors, devoting concentrated periods of time to academic work and reading related but unassigned work." Based on equally weighted class means the level of participation in academic life had increased significantly ($p < .001$) among the undergraduates. The second scale, titled Intellectual Orientation, purports to measure the disposition of an individual toward the creation, development and application of new ideas and the preference for independent thought. In this case neither the difference between the weighted means nor those between the class means for the two surveys were significant.

SUMMARY

Generally, the concern of institutional research has dealt with the economic and administrative aspects of the university. We are advocating the systematic study of the educational effects of intervention activities as an equally important research function. Additionally it is our contention that the procedure we have employed is viable during this period of tight budgets.

While the results to date are encouraging, we are very much aware of the pitfalls in attempting to attribute the positive gains directly to the formalized program of innovation. While studies of this type are not rigorous in the experimental sense and quite susceptible to criticism, they are necessary and valuable. To be sure, there are difficult technical problems. For example, how valid is the notion of perceived actual progress as a proxy measure for the more direct assessment of the outcomes of higher education? No less important is the problem of faculty acceptance. In the original evaluations of the projects it was virtually

impossible to have faculty talk in terms of student outcomes. The seemingly unfamiliarity with the notion along with a distrust of behavioral measurement techniques were major contributing factors. To effectively deal with such problems and to bring about the necessary methodological refinements greater experience with longitudinal data of the type presented here, along with the subsequent dissemination and discussion of its implications, is needed.

NOTES

[1]See "Innovation at Illinois State University" ERIC number: ED082694. (A limited number of copies are available from the first author.)

[2]The statements were assembled from a variety of sources; however, the major source was the Higher Education Measurement & Evaluation Kit developed under the directorship of C. Robert Pace at UCLA, Center for the Study of Evaluation.

[3]KIT refers to the Higher Education Measurement & Evaluation Kit.